

Chi-Square Test

Inference tests for multiple proportions

Questions to Answer

- Do test results support Mendel's genetic principles?

Questions to Answer

- Do test results support Mendel's genetic principles?
- Was surviving the Titanic sinking independent of a passenger's status?

Questions to Answer

- Do test results support Mendel's genetic principles?
- Was surviving the Titanic sinking independent of a passenger's status?
- Do students, teachers and staff show the same distributions in types of cars driven?

Questions to Answer

- Do test results support Mendel's genetic principles? (goodness-of-fit)
- Was surviving the Titanic sinking independent of a passenger's status?
- Do students, teachers and staff show the same distributions in types of cars driven?

Questions to Answer

- Do test results support Mendel's genetic principles? (goodness-of-fit)
- Was surviving the Titanic sinking independent of a passenger's status? (independence)
- Do students, teachers and staff show the same distributions in types of cars driven?

Questions to Answer

- Do test results support Mendel's genetic principles? (goodness-of-fit)
- Was surviving the Titanic sinking independent of a passenger's status? (independence)
- Do students, teachers and staff show the same distributions in types of cars driven? (homogeneity)

Goodness-of-Fit

Ask the question...

- Does this new set of data that we have gathered match (or fit) with these existing results?

Example

The Graying of America

- Better medicine and healthier lifestyles are extending life spans.
- We have data from 1980 and are comparing our collected data from 1996.

US Population by age group (1980)

Age Group	Population (in thousands)	Percent
0 to 24	93,777	41.39
25 to 44	62,716	27.68
45 to 64	44,503	19.64
65 & older	25,550	11.28
Total	226,546	100.00

Hypothesis

H_0 :

H_a :

Hypothesis

Ho: The age group distribution in 1996 is the same as in 1980

Ha: The age group distribution in 1996 is different from 1980

SRS from 1996

Age Group	Population
0 to 24	177
25 to 44	158
45 to 64	101
65 & older	64
Total	500

SRS from 1996

Age Group	Population	Percent
0 to 24	177	35.4
25 to 44	158	31.6
45 to 64	101	20.2
65 & older	64	12.8
Total	500	100.0

Expected Counts

Age Group	1980 Percentages	1996 Expected Counts
0 to 24	41.39	$500(0.4139) = 207.0$
25 to 44	27.68	
45 to 64	19.64	
65 & older	11.28	
	100.00	

Expected Counts

Age Group	1980 Percentages	1996 Expected Counts
0 to 24	41.39	$500(0.4139) = 207.0$
25 to 44	27.68	$500(0.2768) = 138.4$
45 to 64	19.64	$500(0.1964) = 98.2$
65 & older	11.28	$500(0.1128) = 56.4$
	100.00	500

Goodness-of-Fit

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

O = Observed values

E = Expected values

Goodness-of-Fit

Age Group	Observed “O”	Expected “E”	$(O - E)^2/E$
0 to 24			
25 to 44			
45 to 64			
65 & older			

Goodness-of-Fit

Age Group	Observed “O”	Expected “E”	$(O - E)^2/E$
0 to 24	177	207.0	
25 to 44	158	138.4	
45 to 64	101	98.2	
65 & older	64	56.4	

Goodness-of-Fit

Age Group	Observed “O”	Expected “E”	$(O - E)^2/E$
0 to 24	177	207.0	4.3478
25 to 44	158	138.4	2.7757
45 to 64	101	98.2	.0798
65 & older	64	56.4	1.0241
			8.2275

Conclusion

- Degrees of freedom ($n - 1$): $4 - 1 = 3$
- Chi-square test statistic = 8.227
- Falls between critical values (Table E) of .05 and .025
- Reject null hypothesis

Density Curves

- Total area equal to 1
- Begins at 0 (horizontal), reaches peak and approaches horizontal asymptotically
- Skewed right and becomes more symmetric as df increases

Follow Up Analysis

- Use results from goodness-of-fit test
- Be sure to look at individual components for full story

You Try...

- P.716-717 #'s 13.10 and 13.12